Two-Color Data Processing

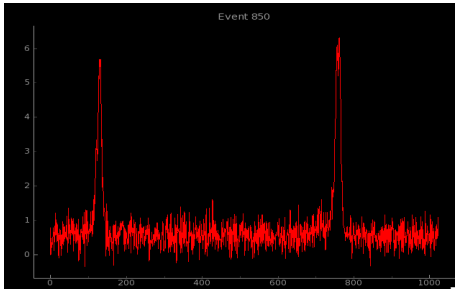
DIALS 6

May 27, 2015

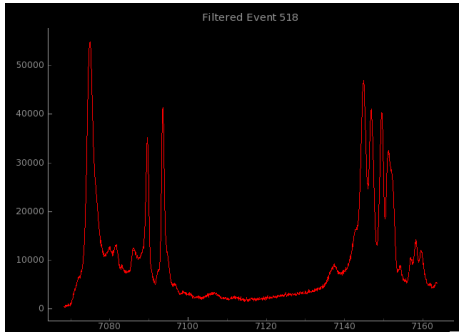
Tara Michels-Clark

Introduction

- XFEL pulses separated in energy and time
  - Self seed pulses may be used for experimental phasing
  - SASE two color pulses at different time separations used to study radiation damage



Self-seeded 90eV separated shot



SASE 60eV separated shot

Introduction

Two Color Lysozyme Diffraction

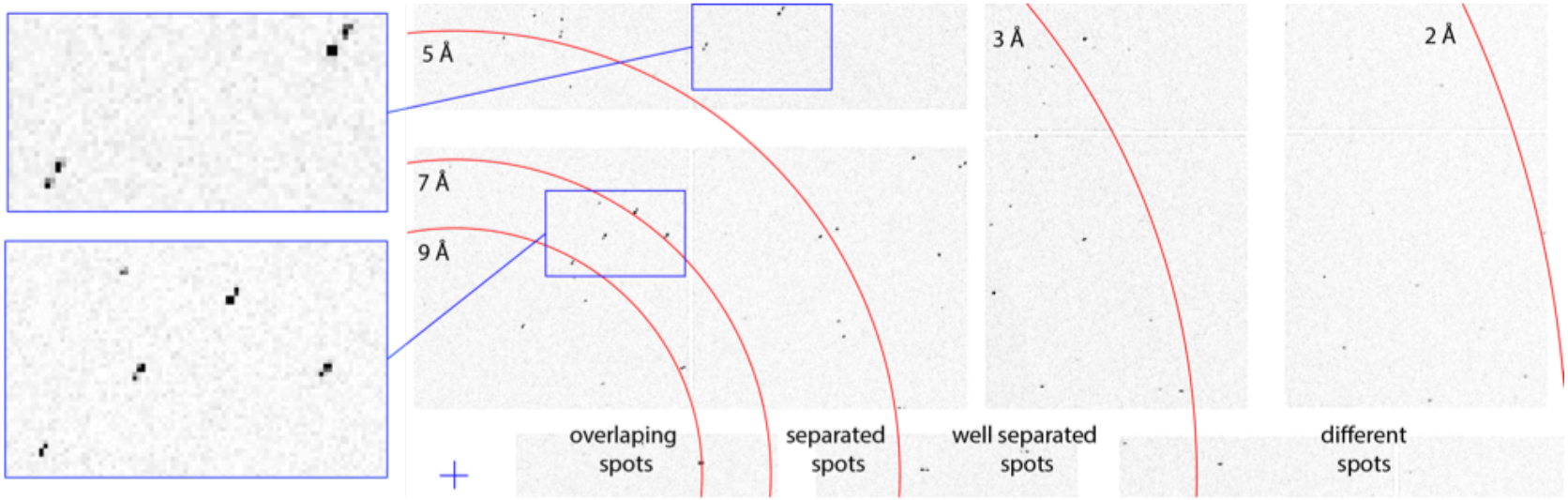


Image courtesy of Soichi Wakatsuki

Processing the Measured Two Color Data

General Protocol:

- Determine two color shots using FEE (Front End Enclosure) spectrometer
- Match the time stamps to images (mod_spectrum_filter)
- run Diffraction filter (hit_finder)
- Index the two color hits
- Predict spots and integrate

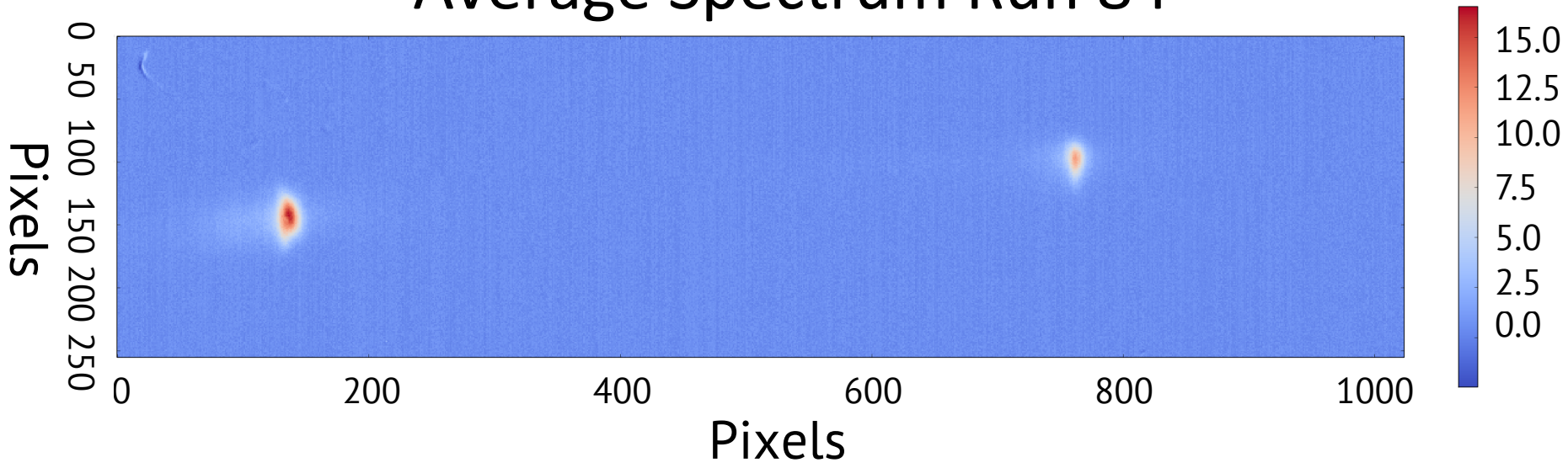
Filtering using FEE

Self Seeded Two Color Filter criteria

- Peak positions
  - Used average peak position over each run
  - Narrow range for low energy peak ±5 pixels
  - Wide range for high energy peak ±100 pixels
- Peak height relative to regions before, after and between peaks
- Peak ratios

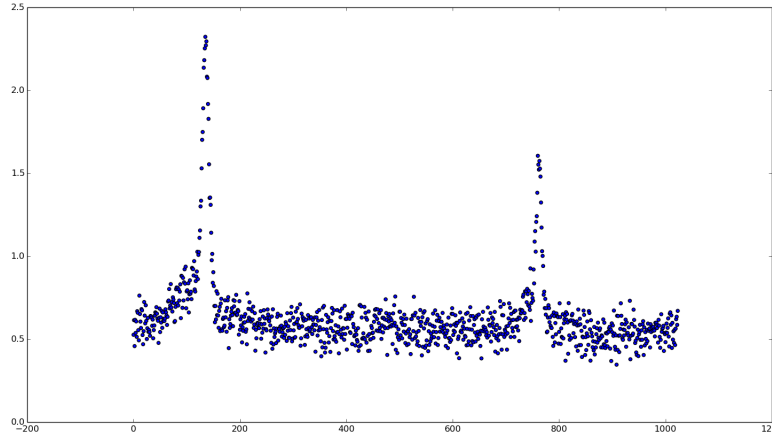
FEE Spectrum

Average Spectrum Run 84



Low energy peak position
~130 pixels

High energy peak position
~760 pixels



Indexing Overview

- Three main parts:
  1. Calculate reciprocal lattice vectors from spot centroid positions
  2. Analyze reciprocal lattice vectors for periodicity
     - Shortest vectors are binned and analyzed for a unique basis
  3. Best candidate crystal matrix (most consistent with centroid positions) and reciprocal lattice vectors are used to calculate *hkl* values

$$\mathbf{h} = \mathbf{A}^{-1}\mathbf{r}$$

Indexing Spots from Two Wavelengths

- FFT algorithm based in real space* to calculate candidate basis using two wavelength functional

$$F_\lambda(\mathbf{x}(\psi,\theta)) = \sum_j \cos(2\pi \mathbf{r}_j(\lambda_1) \bullet ||\mathbf{x}||\hat{\mathbf{u}}(\psi,\theta)) + \sum_j \cos(2\pi \mathbf{r}_j(\lambda_2) \bullet ||\mathbf{x}||\hat{\mathbf{u}}(\psi,\theta))$$

- Calculate fractional *hkl*s from rlps

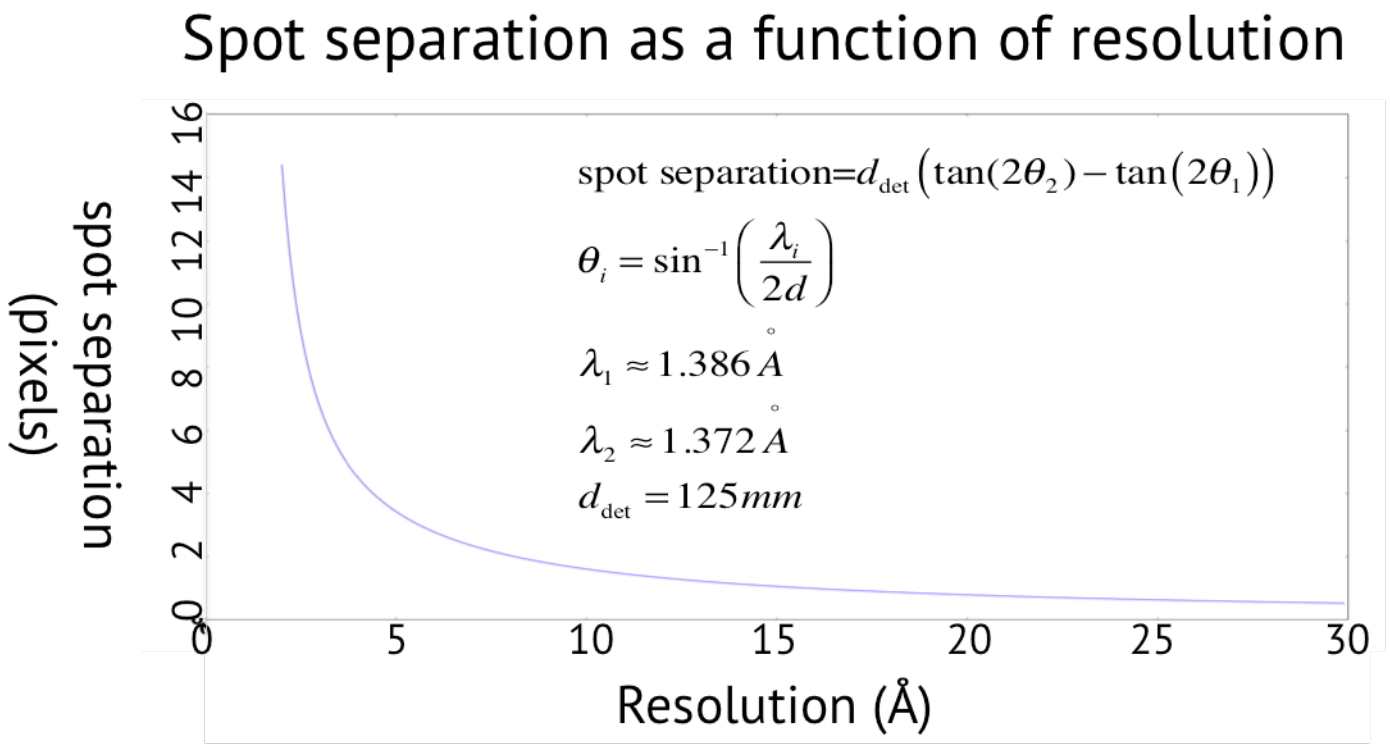
- Rlp with smallest difference norm from integer hkl is assigned to the corresponding wavelength

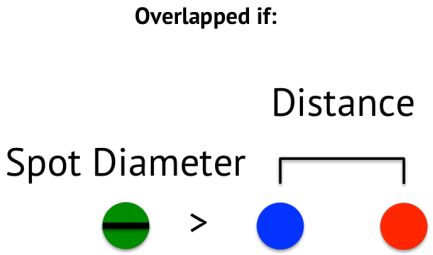
- If more than one spot at the same wavelength can be assigned to the same *hkl* → assign it to the closest one.

* Gildea, R. J., Waterman, D. G., Parkhurst, J. M., Axford, D., Sutton, G., Stuart, D. I., Sauter, N. K., Evans, G. & Winter, G. (2014). Acta Cryst. D70, 2652-2666.

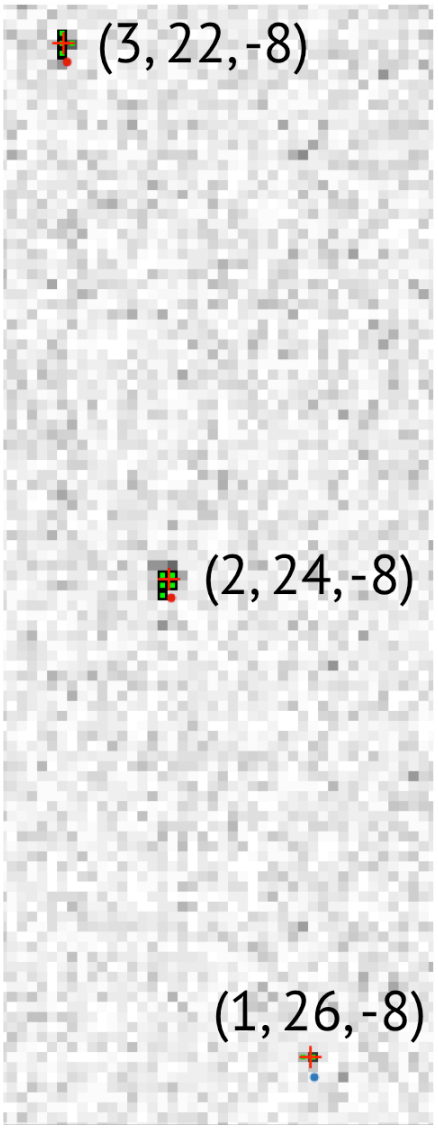
Indexing Spots from Two Wavelengths

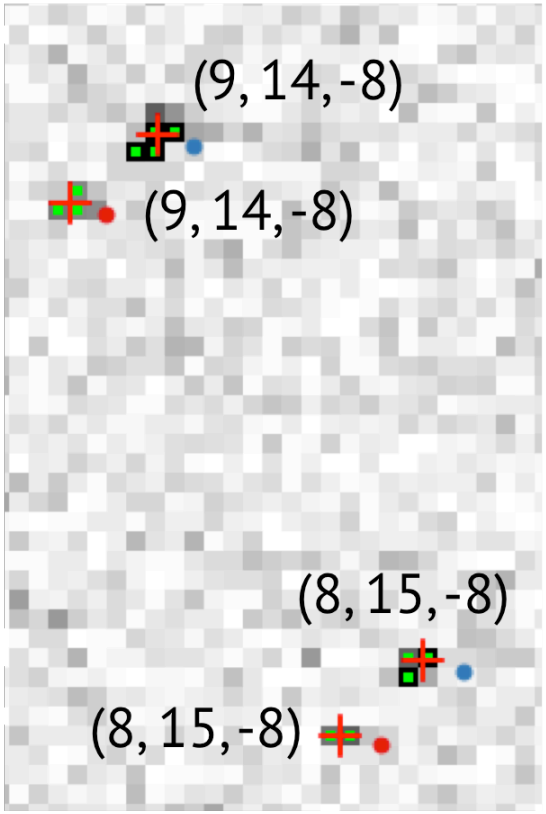
- Overlapped spots in the low resolution region are assigned to the average wavelength

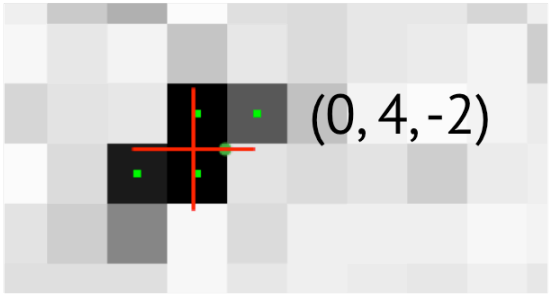




Indexing Results







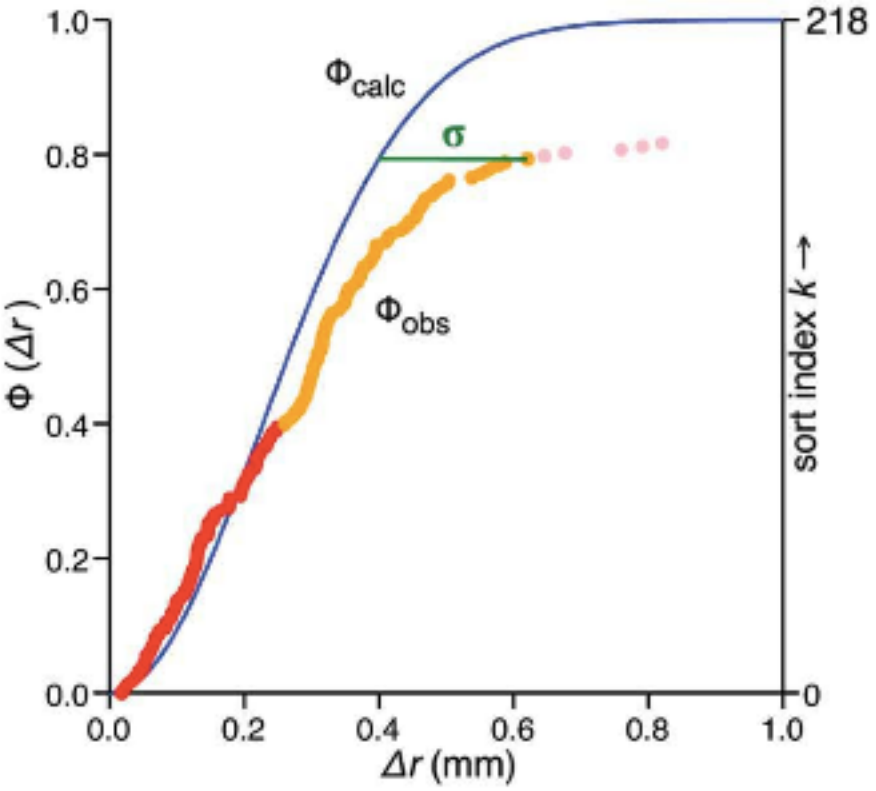
- Spot Centroid
- Valid Pixel (inside spot)
- Overlapped Spot
- Low Energy Spot (or $\lambda_1 \approx 1.386$ Å)
- High Energy Spot (or $\lambda_2 \approx 1.372$ Å)

Integration

- Select the best candidate orientation matrix based on RMSD and area under the green curve
  - Reject outliers based on Sauter and Poon (2010) outlier rejection algorithm
  - Reindex
  - Delta psi outlier rejection
  - Select orientation with smallest green curve volume
- Integrate each experiment separately using same crystal model for each

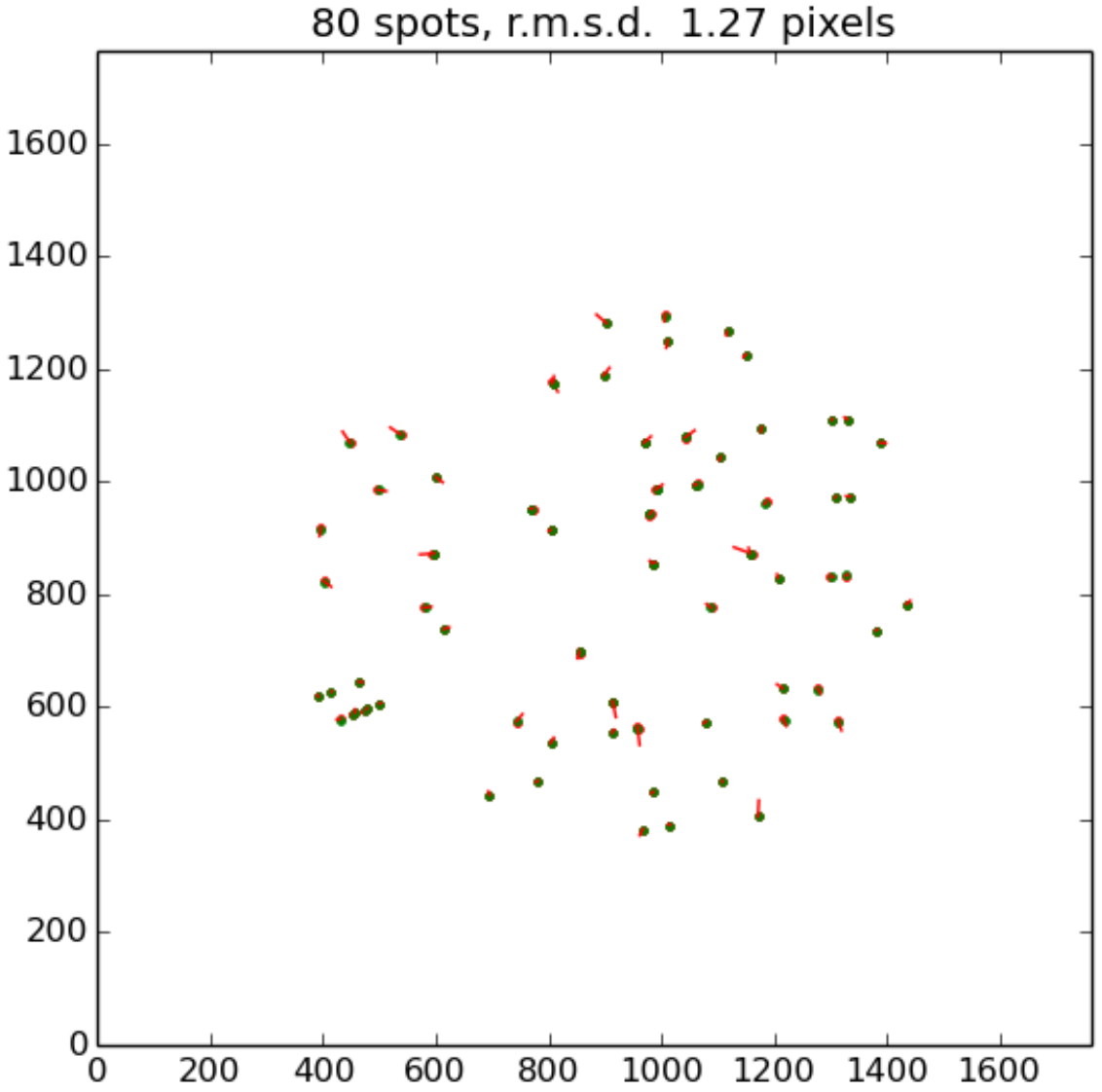
RMSD Outlier Rejection

spots with the largest deviation from the expected cumulative Rayleigh probability distribution are rejected

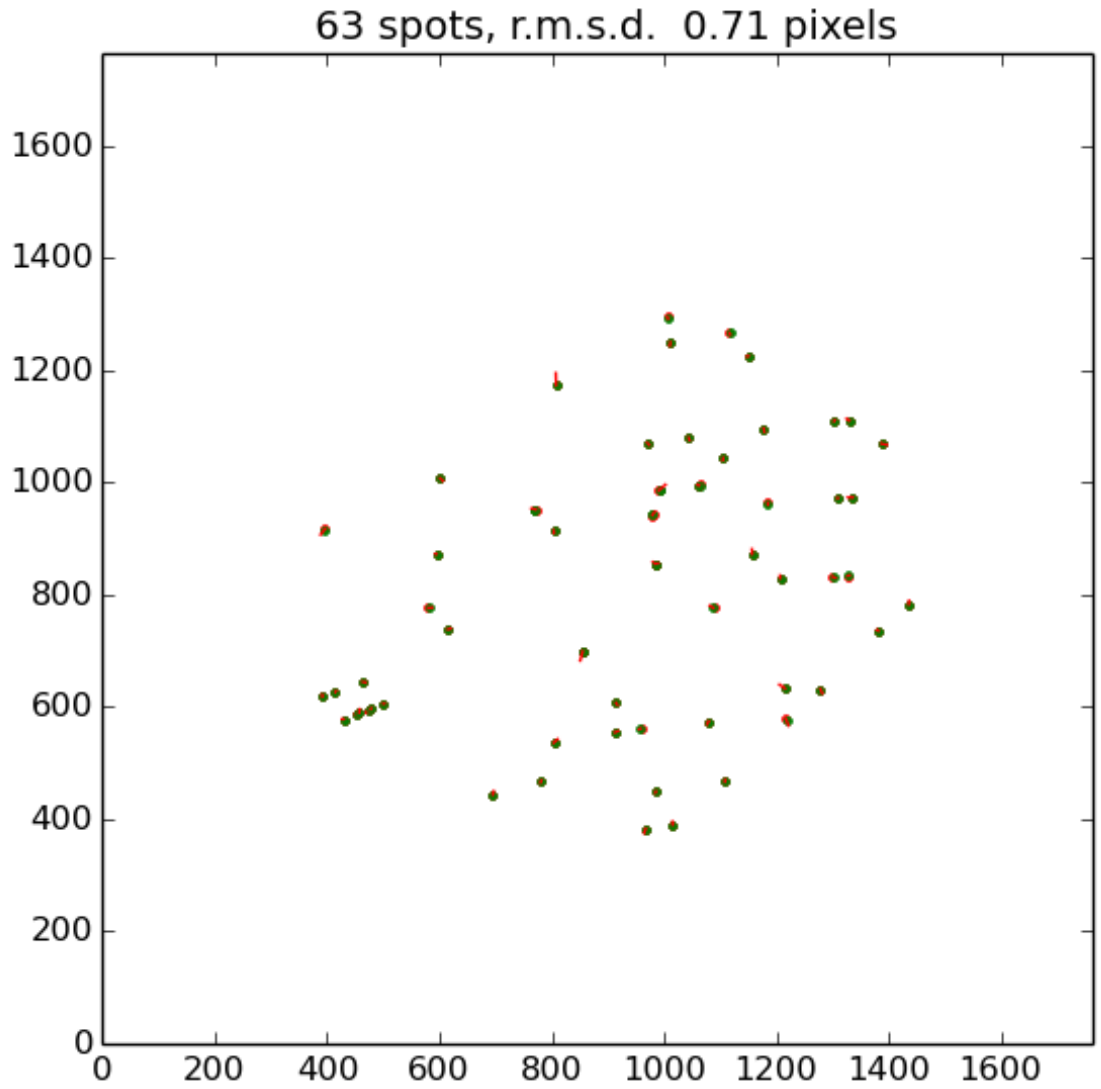


Sauter, N. K. and Poon, B. K. (2010). J. Appl. Cryst.**43**, 611-616.

Initial RMSD and Displacement Vectors



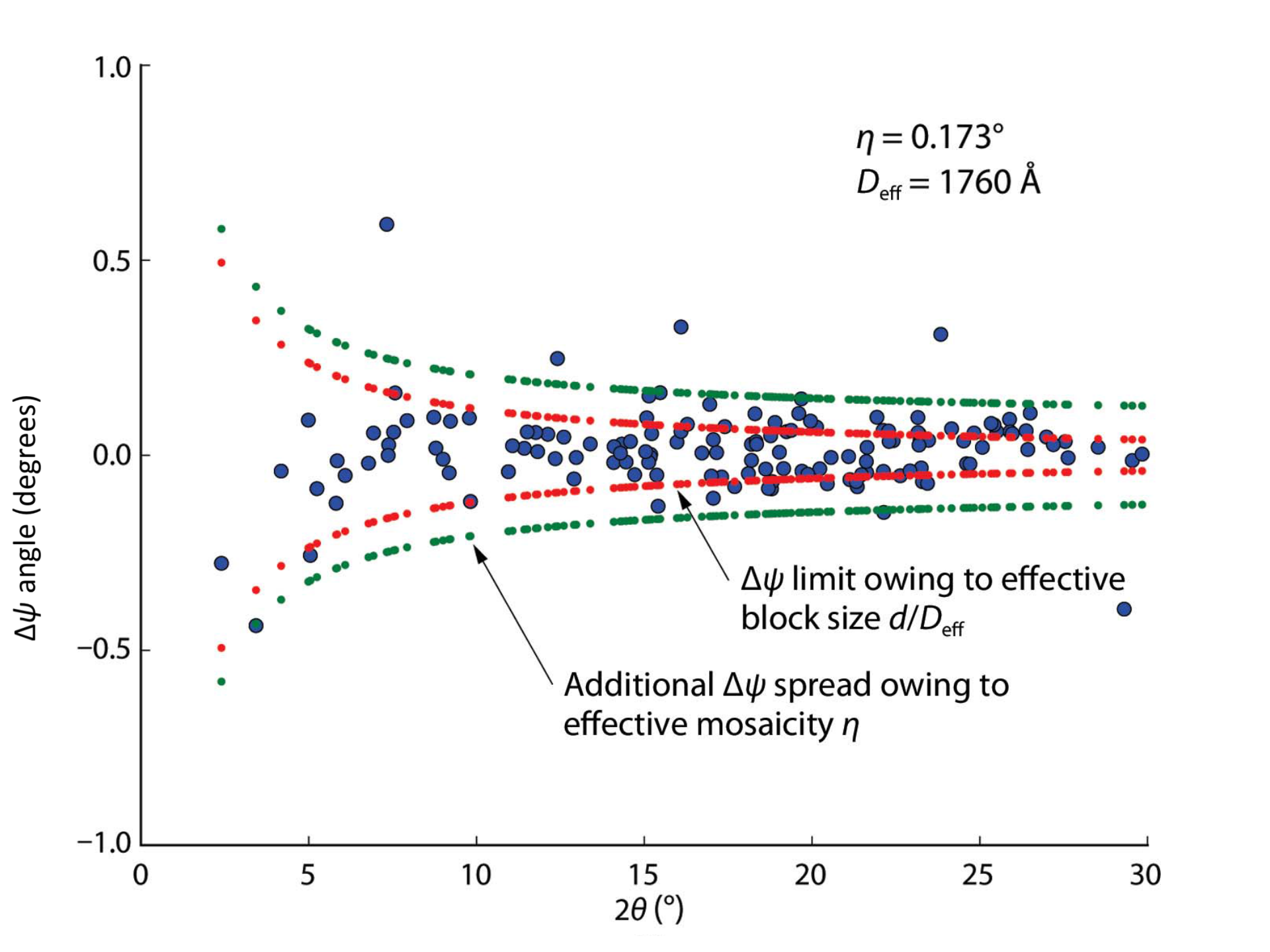
After RMSD Outlier Rejection



Further Optimizing Crystal Model

- Nave parameters* based on the crystal model are calculated
  - Effective domain size D_{eff}
  - Mosaicity η
- Based on these calculations the Δψ limits are calculated

*Nave C (2014). "Matching X-ray beam and detector properties to protein crystals of different perfection." *Journal of synchrotron radiation* **21**

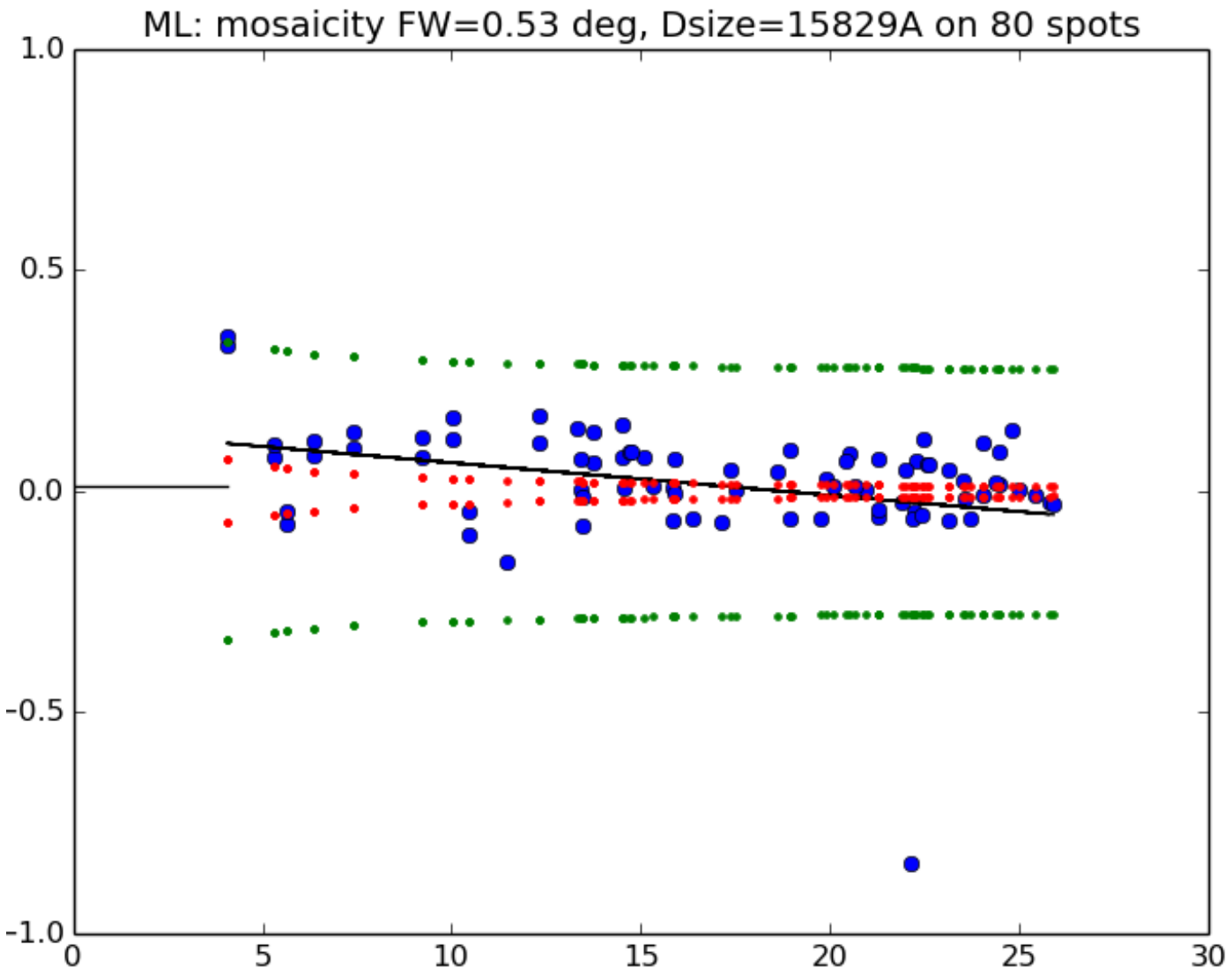

1.0
0.5
0.0
−0.5
−1.0
0
5
10
15
20
25
30
Δψ angle (degrees)
2θ (°)
η = 0.173°
D_eff = 1760 Å
Δψ limit owing to effective block size d/D_eff
Additional Δψ spread owing to effective mosaicity η

Δψ Crystal Model Optimization

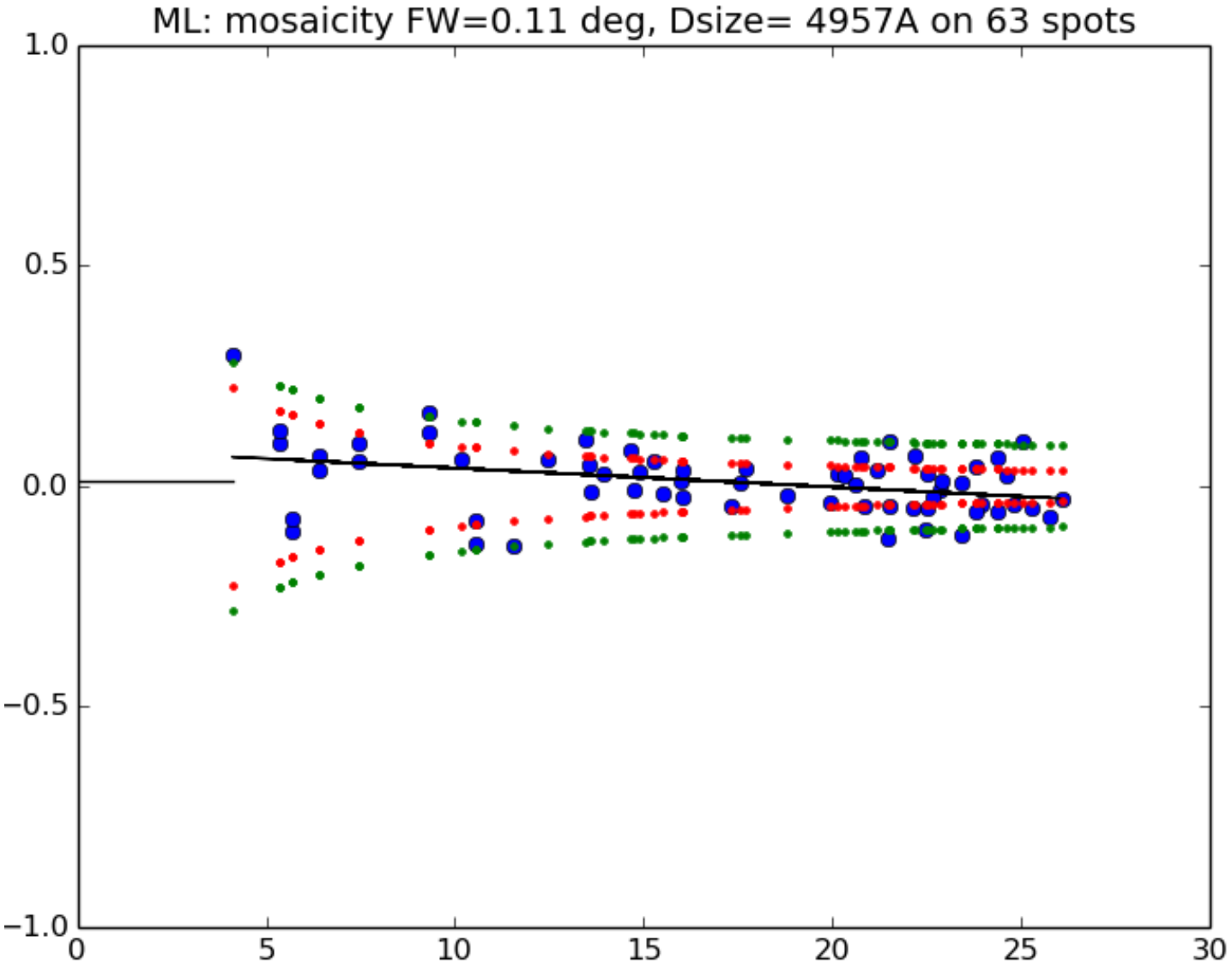
- Spots outside of Δψ* limit are rejected
  - Remaining spots are reindexed using candidate orientation matrix (matrices)
  - The candidate orientation matrix yielding the smallest green curve volume (Δψ spread due to larger effective mosaicity) is retained.

*Sauter NK, Hattne J, Brewster AS, Echols N, Zwart PH, Adams PD (1 Dec 2014): "Improved crystal orientation and physical properties from single-shot XFEL stills." Acta Crystallogr. D Biol. Crystallogr. 70, 3299-309

Δψ Plot Prior to Outlier Rejection



Δψ Plot After Outlier Rejection



Future Work

- Reprocessing both data set with best filter parameters
- Analyze statistics of integration results
- Spot deconvolution for overlapped spots in low resolution region

Acknowledgements

Berkeley National Lab
Nicholas Sauter
Muhamed Amin
Tara Michels-Clark
Iris Young
Nat Echols
Paul Adams
Peter Zwart
Vittal Yachandra
Junko Yano
Jan Kern
James Holton

Janelia Farm
Johan Hattne

LCLS
Uwe Bergmann
Alberto Lutman
…and many others

Diamond Light Source
David Stuart
Gwyndaf Evans
Graeme Winter
Jonathan Grimes
Richard Gildea
James Parkhurst
Luis Fuentes-Montero

CCP4
- David Waterman

UCLA
David Eisenberg
Duilio Cascio
Michael Sawaya
Jose Rodriguez
Luki Goldschmidt

Stanford School of Medicine
Soichi Wakatsuki
Ulf Lundström

IBS
Jacques-Philippe Colletier

Stanford University
Axel Brunger
Mona Uervirojnangkoorn
Artem Lyubimov
Oliver Zeldin

SSRL
Mike Soltis
Ana Gonzalez
Ashley Deacon
Aina Cohen
Yingssu Tsai
Scott McPhillips

BNL
Allen Orville

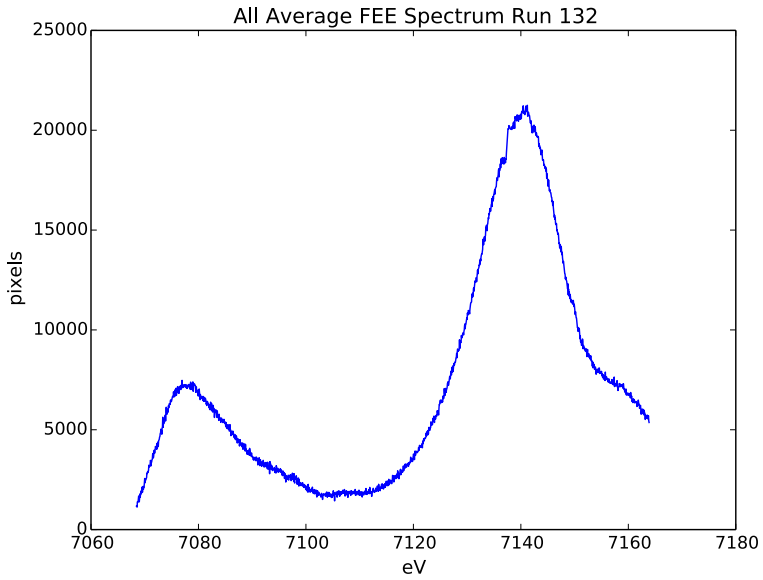
NIH/NIGMS grants 1R01GM095887 and 1R01GM102520
DOE/Office of Science contract DE-AC02-05CH11231

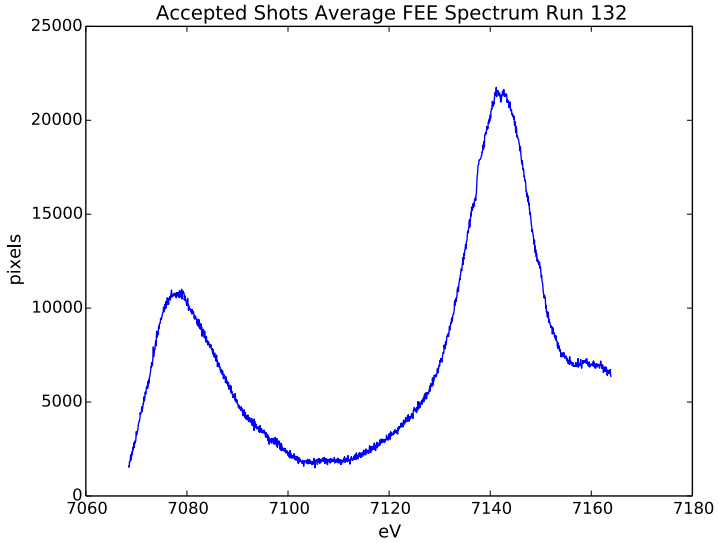
Filtering using FEE

SASE Two Color Filter criteria

- Peak positions
  - Peaks were restricted to be ±10 eV from iron edge (7112 eV)
- Peak ratios
  - Restricted to be 0.10

FEE SASE Spectrum





Run 132

35053 events

13581 rejected